Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 1 February 2008 (MN LATEX style file v1.4)

Higher Order Clustering in the Durham/UKST and Stromlo-APM Galaxy Redshift Surveys

Fiona Hoyle[1], Istvan Szapudi[1,2], Carlton M. Baugh[1]
1. Department of Physics, Science Laboratories, South Road, Durham DH1 3LE
2. Present Address: CITA, University of Toronto, 60 George Street, Ontario, Canada, M5S 3H8,

1 February 2008

ABSTRACT

We present a counts-in-cells analysis of clustering in the optically selected Durham/UKST and Stromlo-APM Galaxy Redshift Surveys. Minimum variance estimates of the second moment, skewness (S_3) and kurtosis (S_4) of the count probability distribution are extracted from a series of volume limited samples of varying radial depth. The corresponding theoretical error calculation takes into account all sources of statistical error on the measurement of the moments, and is in good agreement with the dispersion over mock redshift catalogues. The errors that we find on S_3 and S_4 are larger than those quoted in previous studies, in spite of the fact that the surveys we consider cover larger volumes. S_3 varies little with cell size, with values in the range $1.8-2.2$ and errors $\lesssim 20\%$, for cubical cells of side $3-20h^{-1}$Mpc. Direct measurements of S_3 are possible out to $\sim 35h^{-1}$Mpc, though with larger errors. A significant determination of S_4 is only possible for one scale, $l \sim 6h^{-1}$Mpc, with $S_4 \approx 5$. We compare our results with theoretical predictions from N-body simulations of cold dark matter universes. Qualitatively, the skewness of the dark matter has the same form as that of the galaxies. However, the amplitude of the galaxy S_3 is lower than that predicted for the dark matter. Our measurements of S_3 are consistent with the predictions of a simple model in which initially Gaussian fluctuations in the dark matter evolve gravitationally, if a second order bias term is specified, in addition to the traditional linear bias, in order to describe the relation between the distribution of galaxies and dark matter.

Key words: methods: numerical - methods: statistical - galaxies: formation - large-scale structure of Universe

1 INTRODUCTION

Maps of the local universe have improved dramatically over the last decade and permit the clustering pattern of galaxies to be quantified on large scales (e.g. Efstathiou et al. 1990(a); Maddox et al. 1990; Saunders et al. 1991). Such observations can potentially constrain both the nature of the dark matter and the statistics of primordial density fluctuations.

The first accurate measurements of the galaxy two-point correlation function on scales greater than $10h^{-1}$Mpc indicated more structure than expected in the simplest form of the cold dark matter (CDM) model. This led to variants of the CDM model being studied (Efstathiou, Sutherland & Maddox 1990). Currently, the most successful CDM model is a low density, spatially flat universe with a cosmological constant, ΛCDM. The power spectrum in the ΛCDM model is described by a shape parameter $\Gamma = 0.2-0.3$ (in this Letter we use the parameterisation of the power spectrum given in Efstathiou, Bond & White 1992). If fluctuations in the dark matter are normalised so as to reproduce the local abundance of hot X-ray clusters (White, Efstathiou & Frenk 1993), the power spectrum in the ΛCDM model is similar to that observed for galaxies on scales around $k \sim 0.05 - 0.2h$Mpc^{-1} (Gaztañaga & Baugh 1998). On small scales, however, when the effects of peculiar velocities are ignored (real space), the dark matter power spectrum has a higher amplitude than the galaxy power spectrum (Gaztañaga 1995; Peacock 1997; Jenkins et al. 1998). Furthermore, the small scale power spectrum for galaxies is a power law over a decade and a half in wavenumber, whereas the dark matter correlation function shows considerable curvature.

Heuristic biasing schemes, in which the galaxy distribution is proposed to be a local transformation of the smoothed density field, have enjoyed a certain degree of success in reproducing the observed correlation function (Coles 1993; Cole et al. 1998; Mann, Peacock & Heavens 1998; Narayanan et al. 1999). Progress towards a physical understanding of the processes responsible for producing a bias between the galaxy and dark matter distributions has been made using semi-analytic models for galaxy formation (Benson et al. 2000a,b; Kauffmann et al. 1999). In a ΛCDM model that reproduces the bright end of the field galaxy luminosity function, Benson

Survey	cell size (h^{-1}Mpc)	R_{max} (h^{-1}Mpc)	Volume ($10^6 h^{-3}$Mpc^3)	M_{crit} - 5log$h$	N_{gal}	S_3	S_4
Durham/UKST	3.125	170	0.721	-19.58	510	1.94±0.14	1.5
Durham/UKST	6.3125	180	0.855	-19.73	515	2.11±0.08	5.0±3.8
Durham/UKST	12.625	180	0.855	-19.73	515	1.82±0.21	3.0
Durham/UKST	25.	170	0.721	-19.58	510	1.67±1.32	2.2
Stromlo-APM	3.9375	180	2.547	-19.45	471	2.07±0.57	13.
Stromlo-APM	8.875	180	2.547	-19.45	471	1.89±0.17	3.1
Stromlo-APM	18.1875	190	2.995	-19.58	465	2.24±0.29	8.2
Stromlo-APM	36.625	200	3.493	-19.71	434	1.41±1.01	-

Table 1. Minimum variance estimates of S_3 and S_4 in cubical cells from the Durham/UKST and the Stromlo-APM Surveys. The errors on S_3 are the 1σ theoretical errors for a sample with the volume, geometry and number of galaxies used in the measurement. The relative errors on the estimates of S_4 are greater than 100 per cent apart from one Durham/UKST value.

et al. find remarkably good agreement with both the amplitude and power law slope of the correlation function of APM Survey galaxies (Baugh 1996). If the distortions to the clustering pattern caused by peculiar motions are included, the correlation function of the dark matter is very similar to that of the semi-analytic galaxies in the ΛCDM model, with no bias seen on small scales. The correlation function is also in good agreement with the measurements from galaxy redshift surveys (cf Fig. 1 of Benson et al 2000b).

The constraints on models of galaxy formation provided by the two-point correlation function are somewhat limited. The second moment gives a full statistical description of the density field only in the case of very weak fluctuations. Galaxy clustering can be described in more detail if the J-point, volume-averaged, correlation functions, $\bar{\xi}_J$, are extracted. If the clustering results from the gravitational amplification of a Gaussian primordial density field, then the J-point functions are predicted to follow a hierarchical scaling, $\bar{\xi}_J = S_J \bar{\xi}_2^{J-1}$. The amplitudes S_J do vary with scale, but at a much slower rate than the volume-averaged correlation functions (Juszkiewicz, Bouchet & Colombi 1993; Bernardeau 1994). This scaling behaviour has been studied extensively for cold dark matter in N-body simulations (e.g. Bouchet, Schaeffer & Davis 1991; Baugh, Gaztañaga & Efstathiou 1995; Gaztañaga & Baugh 1995; Hivon et al. 1995; Colombi et al. 1996; Szapudi et al. 1999b).

Fry & Gaztañaga (1993) proposed a simple bias model, based on the assumption that fluctuations in the galaxy distribution can be written as a function of the dark matter fluctuations, when both fields are smoothed on large scales where $\bar{\xi}_2 \ll 1$. The model gives predictions for the moments of the galaxy distribution in terms of the moments for the dark matter. To leading order in the dark matter variance, the galaxy variance is given by $\bar{\xi}_2^{\rm gal} = b^2 \bar{\xi}^{\rm DM}$, where b is usually called the linear bias. To the same order, an additional or second order bias factor, b_2, is required to specify the galaxy skewness:

$$S_3^{\rm gal} = \frac{1}{b}(S_3^{\rm DM} + 3\frac{b_2}{b}). \tag{1}$$

Gaztañaga & Frieman (1994) discuss the implications of the measurements of S_J from the APM Survey for the bias parameters in this model.

In this Letter, we analyse the clustering in two optically selected redshift surveys that sample large volumes of the local universe. The Durham/UKST Survey (Ratcliffe et al. 1998) and Stromlo-APM Survey (Loveday et al. 1996) are magnitude limited to $b_J \approx 17$. Galaxies are sparsely sampled from the parent catalogues at a rate of 1-in-3 in the case of the Durham/UKST Survey and 1-in-20 for the Stromlo-APM Survey. The Stromlo-APM Survey covers a three times larger solid angle than the Durham/UKST Survey. By combining the results from the two surveys, the S_J can be determined over a large dynamic range in cell size.

2 COUNTS-IN-CELLS METHODOLOGY

The technique of measuring the distribution of galaxy counts in cells is well established as a means of quantifying large scale structure (Peebles 1980). The basis of the method is to throw a large number of cells onto the galaxy distribution in order to obtain the probability distribution of finding N galaxies in a cell of a given size l. The moments of the count probability distribution are estimated using the factorial moment technique, which automatically adjusts the moments to compensate for the sampling of a continuous density field using discrete galaxies (Szapudi & Szalay 1993; Szapudi, Meiskin & Nichol 1996). The approach that we adopt here differs in two respects from most previous work. A similar methodology is applied to the PSCz Survey by Szapudi et al. (2000).

The first difference lies in how the higher order moments are extracted from the redshift survey. The count probability distribution is measured in a series of volume limited samples of varying radial depth drawn from the flux limited survey. The moments obtained for a particular cell volume are compared between the different volume limited samples and the minimum variance estimate is adopted as our measurement for this scale. The construction of volume limited samples is straightforward: a maximum redshift for the sample is defined and any galaxy from the flux limited redshift survey that would remain visible if displaced out to this redshift is included in the sample (see, for example, Hoyle et al. 1999).

The number density of galaxies in a volume limited sample is effectively independent of radial distance, with small fluctuations due to large scale structure. This is in direct contrast to a flux limited survey, where the number density changes rapidly with radius. To analyse the count distribution in a flux limited catalogue, a weight must be assigned to each galaxy to compensate for the radial selection function. The analysis of volume limited samples is therefore much simpler, and gives equivalent results without introducing any biases: moreover, the task of devising an optimal weighting scheme and of constructing a suitable estimator of the moments to apply to the flux limited sample is avoided (Colombi, Szapudi & Szalay 1998).

This approach does, however, rely upon the assumption that galaxy clustering does not depend on luminosity, at least over the range of luminosities that we consider in our samples (see column 4 of Table 1 for the absolute magnitudes

that define the volume limited samples we analyse). Loveday et al. (1995) measured the two-point correlation function in redshift space for galaxies selected from the Stromlo-APM survey on the basis of absolute magnitude. These authors found no significant evidence for a difference in clustering amplitude when comparing samples over a much broader range of absolute magnitudes than we consider in our analysis. Similar conclusions were reached by Tadros & Efstathiou (1996) who analysed the amplitude of the power spectrum in different volume limited samples drawn from the same survey. A weak effect, at just over the 1σ level, was seen only for the deepest sample, corresponding to an absolute magnitude of $M_{b_J} = -20.3$. Hoyle et al. (1999) found that the power spectra in volume limited samples drawn from the Durham/UKST survey vary by less than the 1σ errors as the depth of the sample is changed. Therefore, the approximation that the intrinsic clustering in redshift space is the same in different volume limited samples is fully justified by previous work on the surveys we analyse in this Letter.

The second difference from previous work is the treatment of the errors on the measured moments. A theoretical calculation of the errors is made using the method described by Szapudi, Colombi & Bernardeau (1999a) ⋆. All the possible sources of statistical error are included in the calculation, namely the following.

(i) Finite survey volume. The finite volume of the survey means that fluctuations on scales larger than the survey volume are not probed at all. In addition, fluctuations on scales approaching the maximum dimensions of the survey are poorly sampled.
(ii) Edge effects. The density field around galaxies that lie close to the survey boundary is not sampled as well as it is for a galaxy that is well within the boundary. This is because cells are not permitted to straddle the survey boundary.
(iii) Discreteness. The underlying density field is assumed to be continuous. Sampling this field discretely with galaxies makes an additional contribution to the measured moments.
(iv) Sampling or measurement errors due to the finite number of cells used to construct the count probability distribution.

The theoretical calculation of the errors requires a number of quantities to be specified beforehand. Some of these, namely the measured values of the variance and S_J for a given cell size and the sample volume, are estimated directly from the sample. The other quantities, the variance over the full sample volume and the higher order cumulant correlators, are treated as parameters. The errors that we obtain are fairly insensitive to reasonable choices for the values of these parameters (for a full discussion see Szapudi et al. 1999a).

The theoretical error calculation has been extensively tested for clustered distributions of dark matter using N-body simulations (Colombi et al. 2000). As a further check of the calculation, we have compared the results with the dispersion found for the moments averaged over 40 mock Durham/UKST Survey samples with a redshift limit of $z = 0.06$, extracted from the Hubble Volume N-body simulation as described in Hoyle et al. (1999). For each cell size in this comparison, the measurement is selected from the volume limited sample that gives the minimum variance value for the higher order moment. The sample that yielded the best measurement was found to be the same whether the theoretical error or the dispersion over the mock catalogues was used. Furthermore, on this scale, the magnitude of the two error estimates agree to better than 10 per cent. The magnitude of the theoretical errors is within 50 per cent of the dispersion over the mock catalogues on scales that do not give the minimum variance estimates of the higher order moments in a particular sample.

⋆ The FORCE package (FORtran for Cosmic Errors) was used to compute errors. It is available upon request from its authors, S. Colombi (colombi@iap.fr) or IS (szapudi@cita.utoronto.ca).

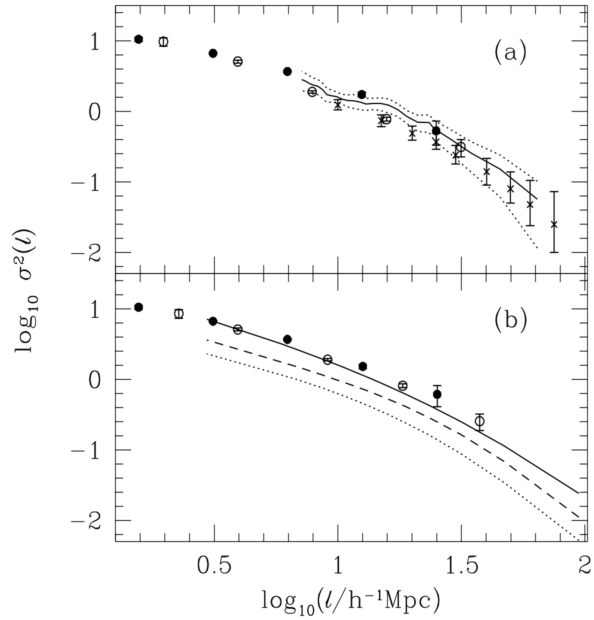

Figure 1. The variance of counts in cubical cells. In both panels, solid circles show the variance in the Durham/UKST Survey, whilst open circles the Stromlo-APM Survey results. In panel (a), we show the variance in volume limited samples with $z_{\max} = 0.06$. The solid line shows an estimate of the variance made from the power spectrum measured in the same Durham/UKST sample by Hoyle et al. (1999); the dotted lines show the 1σ errors. The crosses show the variance for the flux limited Stromlo-APM survey from Loveday et al. (1992). The error bars on the Loveday et al. points show 95 per cent confidence limits. In panel (b), the circles show the best estimates of the variance, extracted from a series of volume limited samples. The lines show the variance in redshift space for the N-body simulations discussed in Section 4: the solid line is for a simulation with a linear power spectrum described by $\Gamma = 0.2$ and $\sigma_8 = 1$, the dashed line for $\Gamma = 0.5$ and $\sigma_8 = 1$ and the dotted line for $\Gamma = 0.5$, $\sigma_8 = 0.66$ (σ_8 is the rms density fluctuation in spheres of radius $8h^{-1}$Mpc).

3 RESULTS

The galaxy count probability distribution is measured in cubical cells of side $3 - 40h^{-1}$Mpc in a series of volume limited samples drawn from the Durham/UKST and Stromlo-APM redshift surveys. The limiting redshifts of the samples are in the range $z \sim 0.05 - 0.08$, corresponding to maximum radial depths of 140–220h^{-1}Mpc. The higher order moments are calculated from the count probability distribution using the factorial moment technique introduced by Szapudi & Szalay (1993). In practice, measurement errors, (iv) in the list of statistical errors given in Section 2, are negligible in comparison to the other contributions, because on the order of 10^8 cells are used to determine the count distribution at each scale. This massive oversampling of the density field is achieved using the algorithm developed by Szapudi (1998).

The second moment or variance of the galaxy distribution is shown in Fig.1. In both panels, the filled circles

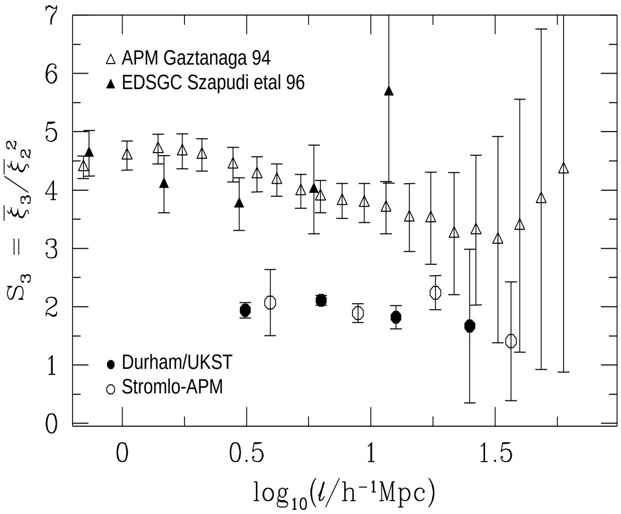

Figure 2. The skewness extracted from the redshift surveys (filled circles show Durham/UKST results, open circles show Stromlo-APM results) compared with the three dimensional values inferred from the parent angular catalogues (the open triangles show the APM Survey results from Gaztañaga 1994, and the filled triangles show the results from the Edinburgh-Durham Southern Galaxy Catalogue from Szapudi, Meiksin & Nichol 1996).

show measurements obtained from the Durham/UKST Survey and the open circles show those from the Stromlo-APM Survey. Fig. 1(a) shows the variance as a function of cell size in volume limited samples extracted from the survey, with a maximum redshift of $z = 0.06$. Fig. 3 of Hoyle et al. (1999) shows that the number of galaxies as a function of the maximum redshift defining a volume limited sample peaks at this redshift for both surveys. These results are in good agreement with estimates of the variance made from the surveys using different techniques. The solid line shows an independent estimate of the variance obtained from the the power spectrum of the same volume limited sample from the Durham/UKST Survey from Hoyle et al. (1999), for wavenumbers $k \leq 0.43h\mathrm{Mpc}^{-1}$. We have used the approximate transformation between power spectrum and variance given in Peacock (1991). The dotted lines show the 1σ error on this estimate, which comes directly from the error on the measured power spectrum. The very good level of agreement between these different estimates demonstrates that large volume cells genuinely measure fluctuations on large scales. Our results for a volume limited subsample of the Stromlo-APM survey agree well with those obtained from the full magnitude limited survey shown by the crosses in Fig. 1(a) (Loveday et al. 1992). The error bars on these points show the 95 per cent percent confidence limits and are computed under the assumption that the distribution of fluctuations is Gaussian.

In Fig. 1(b), the points show the best estimates of the variance extracted from the two surveys as described in Section 2. The best estimates of the variance from the Durham/UKST survey come from two samples, with radial limits of $R_{\max} = 170h^{-1}$Mpc and $R_{\max} = 180\ h^{-1}$Mpc; reading from left to right, the first two points and the last point in Fig. 1(b) come from the $R_{\max} = 170h^{-1}$Mpc sample, whilst the third and fourth points come from the $R_{\max} = 180$ h^{-1}Mpc sample. The smoothness of the locus traced out by the points supports our assumption that there is no significant dependence of clustering strength on luminosity over the samples considered. The lines in 1(b) show the variance in a set of representative CDM simulations; these simulations are discussed in Section 4.

The minimum variance estimates of S_3 from the Durham/UKST and Stromlo-APM surveys are listed in Table 1, along with the properties of the volume limited sample in which the measurement was made. The errors on S_3 are the 1σ theoretical errors predicted for a sample of this volume and geometry and containing the stated number of galaxies. For cubical cells between 3 and 20 h^{-1}Mpc, we find remarkably little variation in the value of S_3, with errors in the range $10-20$ per cent, which again provides further evidence against any significant luminosity dependence of clustering. We obtain S_3 on scales larger than $20h^{-1}$Mpc, but with much larger errors.

When the relative error on the estimate of S_J approaches 100 per cent, the perturbative techniques used in the error calculation break down. Nevertheless, the calculation still reliably indicates that the errors are large and that the measurement has no significance. The relative errors on S_4 are estimated to be >100 per cent on all scales in the Stromlo-APM survey. There is only one scale where S_4 can be reliably constrained from the Durham/UKST survey. This scale is also the one for which S_3 is most accurately measured in this sample. As we expect this to be the case in general, the values for S_4 on the same scale as the minimum variance measurements of S_3 are listed in Table 1. These estimates should be treated with caution as the errors are large.

4 DISCUSSION

The mean values we obtain for the skewness are in agreement with those found in shallower redshift surveys, though we find errors that are somewhat larger (e.g. Gaztañaga 1992; Bouchet et al. 1993; Fry & Gaztañaga 1994; Benoist et al 1999 and for a comprehensive compilation of results and a more exhaustive set of references, see table 1 of Hui & Gaztañaga 1999). Moreover, in spite of the relatively large volumes of the surveys considered in this Letter, we find that a significant measurement of S_4 is only possible at one scale. There are two main reasons for the discrepancy in the magnitude of the estimated errors. The first is that some previous results are quoted as averages over the values of S_3 determined on different scales, exploiting the relatively flat form of S_3 in redshift space. This leads to smaller errorbars under the incorrect assumption that the individual measurements are independent. The second reason is that not all of the contributions to the statistical errors listed in Section 2 were considered in previous analyses.

We have constrained S_3 over a wide range of scales, extending beyond $l \sim 20h^{-1}$Mpc, where simple models for bias can be tested most cleanly. Indirect measurements of S_3 on these scales have been obtained from the *IRAS* 1.2-Jy Redshift Survey by fitting a parametric functional form for the count probability distribution to the measured counts (Kim & Strauss 1998). The choice of function is not physically motivated and the error model used is simplistic and may underestimate the true variance (Gaztañaga, Fosalba & Elizalde 1999; Hui & Gaztañaga 1999). Szapudi et al. (2000) have measured S_3 from the IRAS PSCz survey, using the same techniques employed in this paper, and find $S_3 = 0.87 \pm 0.48$ for cells of side $l = 37h^{-1}$Mpc, which is in good agreement with the value we find, quoted in Table 1.

We compare our measurements of S_3 with the values inferred from the parent angular catalogues of the redshift surveys in Fig 2 (Gaztañaga 1994; Szapudi, Meiksin & Nichol 1996). The results from the angular catalogues are obtained by first finding the projected count distribution on the sky,

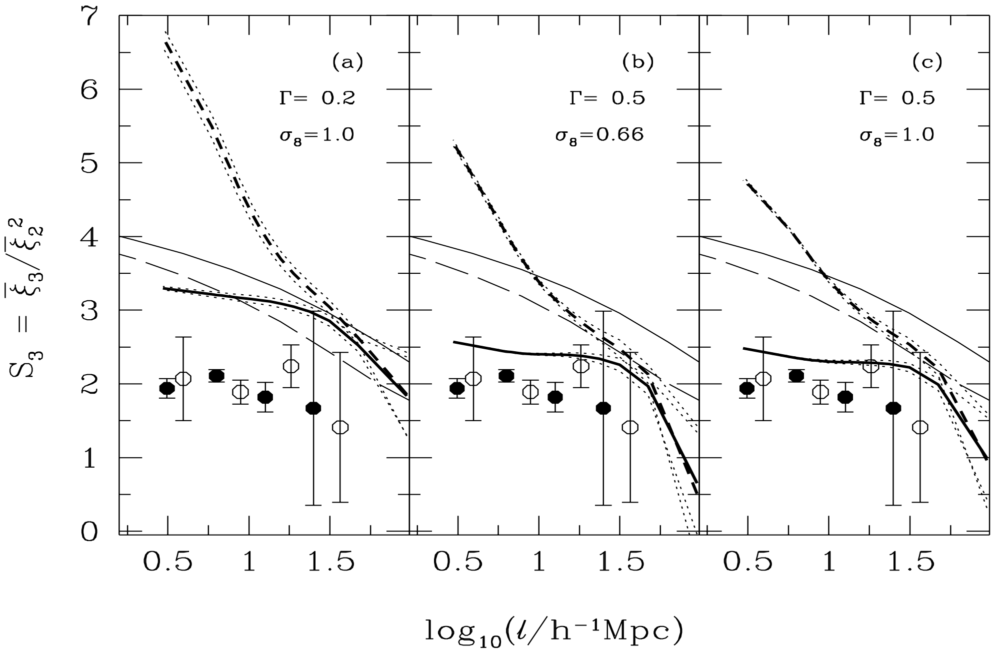

Figure 3. A comparison of the minimum variance measurements of the skewness listed in Table 1 with the skewness obtained from N-body simulations. In each panel, the filled circles show the skewness measured in the Durham/UKST Survey and the open circles show Stromlo-APM Survey results. The light lines show the linear perturbation theory predictions for S_3 in real space and are reproduced in each panel; the solid line shows the skewness for a power spectrum with $\Gamma = 0.2$, and the dashed line shows the result for $\Gamma = 0.5$. The heavy lines show the simulation results and the dotted lines show the error on the mean over five realisations of the initial density field. The heavy dashed (solid) lines show the skewness measured in real (redshift) space. The simulation outputs are described by the following sets of power spectrum parameters: (a) $\Gamma = 0.2$ and $\sigma_8 = 1$, (b) $\Gamma = 0.5$ and $\sigma_8 = 0.66$ and (c) $\Gamma = 0.5$ and $\sigma_8 = 1$.

and then applying a deprojection algorithm to infer the moments in three dimensions. The algorithm requires knowledge of the survey selection function. The deprojected angular measurements are in real space as they are free from any distortion due to the peculiar motions of galaxies. On large scales, the mean value we find for S_3 is below that found in real space. However, the errors are large on both measurements, and the results are consistent at the 1σ level. Moreover, it is somewhat unclear exactly how important edge effects in the angular measurements and systematic effects in the deprojection technique are on these large scales (Szapudi, Meiksin & Nichol 1996; Gaztañaga & Baugh 1998, Gaztañaga & Bernardeau 1998; Szapudi & Gaztañaga 1998).

On small and intermediate scales, $l \leq 15h^{-1}$Mpc, our determinations are below those obtained from the angular catalogues. This is due to redshift space distortions. The same qualitative behaviour is seen for S_3 measured in real space and redshift space in numerical simulations of hierarchical clustering. In Fig. 3, we compare S_3 measured in the N-body simulations used by Gaztañaga & Baugh (1995), which are representative of the behaviour in CDM models, with the redshift survey results. The heavy dashed lines in each panel show S_3 in real space, and the heavy solid lines show S_3 including the effects of the peculiar motions of the dark matter. The dotted lines show the error on the mean obtained over five realisations of the initial conditions (the box size of the simulations is $378h^{-1}$Mpc). Two different power spectra are considered: panel (a) shows a model with $\Gamma = 0.2$ and (b) and (c) show a model with $\Gamma = 0.5$ at two different epochs. On large scales, the value of S_3 depends upon the shape of the power spectrum and is in good agreement with the perturbation theory predictions, which are shown by the light lines; this result was discussed by Gaztañaga & Baugh (1995). The value of S_3 in redshift space also depends upon the shape of the power spectrum, and is insensitive to epoch or equivalently to the amplitude of the fluctuations, as shown by Figs. 3(b) and (c). The real and redshift space values of S_3 become consistent at $l \approx 20h^{-1}$Mpc, in excellent agreement with the comparison presented for the data in Fig. 2.

We now investigate how the predictions from the simulations can be reconciled with the observations and discuss the implications for biasing. The model developed by Fry & Gaztañaga (1993) predicts a relationship between the skewness in the galaxy distribution, $S_3^{\rm gal}$, and that in the underlying dark matter, $S_3^{\rm DM}$, that is applicable on large scales (equation 1). The variance for the dark matter in the simulation with $\Gamma = 0.2$ and $\sigma_8 = 1$ is very close to the observed variance in galaxy counts (cf. the solid line in Fig 1), indicating that a relatively small linear bias term is required; at $l \sim 20h^{-1}$Mpc, the linear bias is $b = 1.16 \pm 0.06$. Furthermore, in redshift space, the linear bias is essentially independent of scale. Thus, given the scale independence of the skewness that we measure for galaxies and which is predicted for the dark matter from the simulations, we can insert the values for $S_3^{\rm gal}$, $S_3^{\rm DM}$ and b into equation 1 and obtain a value for the second order bias term, b_2. At $l \sim 20h^{-1}$Mpc, a second order bias term of value $b_2 = -0.20 \pm 0.14$ is required for the skewness of the dark matter to match that seen for galaxies. For the simulation with $\Gamma = 0.5$ and $\sigma_8 = 0.66$, the linear bias term is larger (cf. the dotted line in Fig. 1), $b = 1.86 \pm 0.10$, and the second order bias term is $b_2 = 1.0 \pm 0.4$. Hence, whilst a linear bias term is sufficient to reconcile the variance measured in redshift space for galaxies and for dark matter, additional bias terms are required to match up the results for the skewness.

A similar counts in cells analysis has been applied to

the PSCz Survey, and yields values for S_3 that are in good agreement with those reported here at all scales (Szapudi et al. 2000). At first sight this result is intriguing, in view of the well known difference in the amplitude of the two-point functions of optical and infra-red selected galaxies on large scales (e.g. Peacock 1997; Hoyle et al. 1999). Thus having demonstrated the need to consider a second order bias term in addition to the linear bias usually discussed, it would appear that both these quantities can depend on the way in which galaxies are selected. These issues are best addressed using semi-analytic models of galaxy formation (Baugh et al. 2000).

ACKNOWLEDGMENTS

FH acknowledges receipt of a PPARC studentship. This work was supported by the PPARC rolling grant at Durham. We thank the Virgo Consortium (see Jenkins et al. 1998 paper) for making the Hubble volume simulation available and for providing software to analyse the output. Tom Shanks and Shaun Cole provided several useful suggestions that improved an earlier draft. We are indebted to the efforts of all those involved in constructing both the parent catalogues and the redshift surveys used here, and for making the redshift catalogues publically available.

REFERENCES

Baugh, C.M., 1996, MNRAS, 280, 267
Baugh, C.M., Gaztañaga, E., Efstathiou, G., 1995, MNRAS, 274, 1049
Baugh, C.M., Szapudi, I., Benson, A.J., Lacey, C.G., Cole, S., Frenk, C.S., 2000, MNRAS, submitted
Benoist, C., Cappi, A., Da Costa, L.N., Mauogordato, S., Bouchet, F.R., Schaeffer, R., 1999, ApJ, 514, 563
Benson, A.J., Baugh, C.M., Cole, S., Frenk, C.S., Lacey, C.G., 2000b, MNRAS, in press
Benson, A.J., Cole, S., Frenk, C.S., Baugh, C.M., Lacey, C.G., 2000a, MNRAS, 311, 793
Bernardeau, F., 1994, A&A, 291, 697
Bouchet, F. R., Schaeffer, R., Davis, M., 1991, ApJ, 383, 19
Bouchet, F. R., Strauss, M.A., Davis, M., Fisher, K.B., Yahil, A., Huchra, J.P., 1993, ApJ, 417, 36
Cole, S., Hatton, S., Weinberg, D.H., Frenk, C.S., 1998, MNRAS, 300, 945.
Coles, P., 1993, MNRAS, 262, 1065
Colombi, S., Bouchet, F.R., Hernquist, L., 1996, ApJ, 465, 14
Colombi, S., Szapudi, I., Jenkins, A., Colberg, J., 2000, MNRAS, 313, 711
Colombi, S., Szapudi, I., Szalay, A. S., 1998, MNRAS, 296, 253
Efstathiou, G., Bond, J. R., White, S. D. M., 1992, MNRAS, 258, 1
Efstathiou, G., Kaiser, N., Saunders, W., Lawrence, A., Rowan-Robinson, M., Ellis, R.S., Frenk, C.S., 1990(a), MNRAS, 247, 10
Efstathiou, G., Sutherland, W.J., Maddox, S.J., 1990(b), Nature, 348, 705
Fry, J. N., Gaztañaga, E., 1993, ApJ, 413, 447
Fry, J. N., Gaztañaga, E., 1994, ApJ, 425, 1
Gaztañaga, E., 1992, ApJ, 398, L17
Gaztañaga, E., 1994, MNRAS, 268, 913
Gaztañaga, E., 1995, ApJ, 454, 561
Gaztañaga, E., Baugh, C.M., 1995, MNRAS, 273, L1
Gaztañaga, E., Baugh, C.M., 1998, MNRAS, 294, 229
Gaztañaga, E., Bernardeau, F., 1998, A&A, 331, 829
Gaztañaga, E., Frieman, J. A., 1994, ApJ, 437, L13
Gaztañaga, E., Fosalba, P., Elizalde, E., 1999, astro-ph/9906296
Hivon, E., Bouchet, F.R., Colombi, S., Juszkiewicz, R., 1995, A&A, 298, 643
Hoyle, F., Baugh, C. M., Shanks, T., Ratcliffe, A., 1999, MNRAS, 309, 659
Hui, L. & Gaztañaga, E., 1999, ApJ, 519, 622
Jenkins, A., et al. 1998, ApJ, 499, 20
Juszkiewicz, R., Bouchet, F., Colombi, S., 1993, ApJ, 412, L9
Kauffmann, G., Colberg, J.M., Diaferio, A., White, S.D.M., 1999, MNRAS, 303, 188
Kim, R. S. J., Strauss, M. A., 1998, ApJ, 493, 39
Loveday, J., Efstathiou, G., Peterson, B. A., Maddox, S. J., 1992, ApJ, 400, L43
Loveday, J., Maddox, S.J., Efstathiou, G., Peterson, B.A., 1995, ApJ, 442, 457
Loveday, J., Peterson, B.A., Maddox, S.J., Efstathiou, G., 1996, ApJS, 107, 201
Maddox, S.J., Efstathiou, G., Sutherland, W.J., Loveday, J., 1990, MNRAS, 242, 43
Mann, R.G., Peacock, J.A., Heavens, A.F., 1998, MNRAS, 293, 209
Narayanan, V., Weinberg, D.H., Branchini, E., Frenk, C.S., Maddox, S., Oliver, S., Rowan-Robinson, M., Saunders, W., 1999, ApJ , submitted, astro-ph/9910229
Peacock, J. A., 1991, MNRAS, 253, 1
Peacock, J. A., 1997, MNRAS, 284, 885
Peebles, P.J.E., 1980, "The Large Scale Structure in the Universe", Princeton University Press, Princeton
Ratcliffe, A., Shanks, T., Parker, Q. A., Broadbent, A., Watson, F. G., Oates, A. P., Collins, A. A., Fong, R., 1998, MNRAS, 300, 417
Saunders, W., Frenk, C.S., Rowan-Robinson, M., Lawrence, A. & Efstathiou, G., 1991, Nature, 349, 32
Szapudi, I., 1998, ApJ, 497, 16
Szapudi, I. & Gaztañaga, E., MNRAS, 1998, 300, 493
Szapudi, I., Meiksin, A. & Nichol, R. C., 1996, ApJ, 473, 15
Szapudi, I., Colombi, S. & Bernardeau, F., 1999a, MNRAS, 310, 428
Szapudi, I., Quinn, T., Stadel, J., Lake, G., 1999b, ApJ, 517, 54
Szapudi, I., Branchini, E., Frenk, C.S. Maddox, S., Saunders, W., 2000, MNRAS, submitted
Szapudi, I. & Szalay, A. S., 1993, ApJ, 408, 43
Tadros, H. & Efstathiou, G., 1996, MNRAS, 282, 1381
White, S.D.M., Efstathiou, G. & Frenk, C.S., 1993, MNRAS, 262, 1023